# Exhibit 24

# INTEGRATION OF USER PROFILES: MODELS AND EXPERIMENTS IN INFORMATION RETRIEVAL

Sung H. Myaeng
4-206 Center for Science & Technology, School of Information Studies,
Syracuse University, Syracuse, NY 13244

and

Robert R. Korfhage*
Department of Information Science, University of Pittsburgh,
Room 737 LIS Building, Pittsburgh, PA 15260

**Abstract** — One difficult problem in information retrieval (IR) is the proper interpreta-
tion of user queries. It is extremely hard for users to express their information needs in
a specific yet exhaustive way. In an effort to alleviate this problem, two theoretical mod-
els have been proposed to utilize user characteristics maintained in the form of a user pro-
file. Although the idea of integrating user profiles into an IR system is intuitively
appealing, and the models seem viable, no research to date has established a foundation
for the roles of user profiles in such a system. Aiming at the investigation of the roles
of user profiles, therefore, this study first identifies and extends various query/profile
interaction models to provide a ground upon which the investigation can be undertaken.
From a continuum of models characterized on the basis of interaction types, metrics, and
parameters, nearly 400 models are chosen to investigate the "model space." New mea-
sures are developed based on the notion of user satisfaction/frustration. In addition,
three different criteria are used to guide users in making judgments on the quality of
retrieved items. Analysis of the data obtained from the experiments shows that, for a
wide variety of criteria and metrics, there are always some query/profile interaction mod-
els that outperform the query alone model. In addition, preferable characteristics for dif-
ferent criteria are identified in terms of interaction types, parameters, and metrics.

## 1. INTRODUCTION

The problem of retrieving information from natural language databases has been studied
during the past quarter century. In traditional context, retrospective information retrieval
(IR) systems are those in which a user initiates the search process by means of a set of ac-
tive queries and receives a set of references to items of potential interest.

One difficult problem in such systems is the transformation of the user's information
need to the form of an explicit query which accurately matches the original intention, and
retrieves all items of interest in the database being searched, and only those. Therefore,
users often have great difficulty in using an IR system successfully regardless of the query
language implementation (e.g., a vector form, a boolean expression of terms, a combina-
tion of both [1,2,3], or other retrieval models [4,5,6,7,8]). As a result, user queries are not
completely satisfactory in expressing the needs in most retrieval situations. It seems nat-
ural that the output of a system based on such a query is necessarily incomplete and
unsatisfactory.

One reason underlying this query formulation problem is the mismatch between terms
used in a query and those used in documents. Blair and Maron [9] analyzed the poor per-
formance of a large-scale IR system and contended that it is exceedingly difficult for users
to predict the exact words, word combinations, and phrases that are used by all (or most)
relevant documents and only (or primarily) by those documents. This difficulty is usually

*All editorial correspondence should be sent to Robert R. Korfhage.